Controlled Markov set-chains and Bayesian Decision Analysis (Decision making theories under uncertainty and its applications : the extensions of mathematics for programming)

Horiguchi, Masayuki

数理解析研究所講究録
第2078巻 2018年 94-101

Controlled Markov set-chains and Bayesian Decision Analysis

Masayuki HORIGUCHI
Professor of Department of Mathematics,
Faculty of Science,
Kanagawa University

Abstract

In this note, we summarize briefly our series of studies on Markov decision processes with unknown transition matrices and give an idea to show the rate of convergence of Bayesian updating of posterior distribution in a sampling problem.

1 Interval estimated Markov decision processes

MDPs consist of four tuples $\{S, A, Q, \boldsymbol{r}\}$ as follows. Let $S = \{1, 2, \ldots, n\}$ be a finite state space, $A = \{a_1, a_2, \ldots, a_k\}$ finite action space. The set of probabilities on S and transition kernels are defined as follows. $P(S) := \{p = (p_1, p_2, \ldots, p_n) \in \mathbb{R}_+^n \big| \sum_{i\in S} p_i = 1\}, P(S|S) := \{q = (q_{ij} : i, j \in S) \in \mathbb{R}_+^{n\times n} \big| \sum_{j\in S} q_{ij} = 1\ (i \in S)\}, P(S|S \times A) := \{Q = (q_{ij}(a) : i, j \in S, a \in A) \in \mathbb{R}_+^{kn\times n} | q_{i\cdot}(a) \in P(S)\ (i \in S, a \in A)\}$, where $\mathbb{R}_+$ is the set of nonnegative real numbers and $\mathbb{R}_+^m$ the set of m-dimensional nonnegative real vectors. Let $B_+(D)$ denote the set of all non-negative real functions on finite set D. For a finite D of n elements , $B_+(D)$ is identified with $\mathbb{R}_+^n$. Let $Q = (q_{ij}(a)) \in P(S|S \times A)$ denote a parameter space of k *unknown transition matrices* and $\boldsymbol{r} = (r(i, a)) \in B_+(S \times A)$ reward function .

For any stationary policy $f \in F$, discounted total expected reward $\phi(f|Q) \in \mathbb{R}_+^n$ with discount factor β $(0 < \beta < 1)$ is defined as a function of stochastic matrix $Q \in P(S|S\times A)$ as:

$$\phi(f|Q) = \sum_{t=0}^{\infty} (\beta Q(f))^t \boldsymbol{r}(f), \tag{1}$$

where, $\boldsymbol{r}(f) = (r(1, f(1)), r(2, f(2)), \ldots, r(n, f(n)))' \in \mathbb{R}_+^n, Q(f) = (q_{ij}(f(i))) \in P(S|S)$.

Q is estimated by interval matrix $\mathcal{Q} = \langle \underline{Q}, \overline{Q}\rangle$, where

$$\begin{aligned} \underline{Q} &= (\underline{q}_{ij}(a) : i, j \in S, a \in A) \in \mathbb{R}_+^{kn\times n}, \\ \overline{Q} &= (\overline{q}_{ij}(a) : i, j \in S, a \in A) \in \mathbb{R}_+^{kn\times n}, \\ \mathcal{Q} &= \langle \underline{Q}, \overline{Q}\rangle = \left\{Q \in P(S|S \times A) \big| \underline{Q} \preceq Q \preceq \overline{Q}\right\}. \end{aligned} \tag{2}$$

It should be noted that the partial orders $\preceq, \prec$ on $\mathbb{R}^{m\times n}$ are defined by the components orders for each element $\underline{q}_{ij}(a)$ and $\overline{q}_{ij}(a)$ $(i, j \in S, a \in A)$. We call $\{S, A, \mathcal{Q}, \boldsymbol{r}\}$ the interval estimated MDPs.

For $f \in F$, we define discounted total expected-set valued value function $\phi(f|\mathcal{Q})$ as follows:

$$\phi(f|\mathcal{Q}) = \left\{\phi(f|Q)\middle|Q \in \mathcal{Q}\right\} \subset \mathbb{R}_+^n \tag{3}$$

where the value $\phi(f|Q)$ of standard MDPs is defined in (1).

Let $C(\mathbb{R}_+)$ denote the set of all bounded and closed intervals in $\mathbb{R}_+$ and $C(\mathbb{R}_+)^n$ the set of all n-dimensional column vectors whose elements are in $C(\mathbb{R}_+)$, i.e.,

$$C(\mathbb{R}_+)^n = \left\{D = (D_1, D_2, \ldots, D_n)'\middle|D_i \in C(\mathbb{R}_+)\ (1 \leqq i \leqq n)\right\}. \tag{4}$$

We will denote by $\mathcal{M}_n$ the set of all interval matrices with $n \times n$ elements.

Lemma 1. *(Hartfiel[3], Kurano, Song, Hosaka and Huang[5])*

(i) Any $\mathcal{Q} \in \mathcal{M}_n$ is a convex polytope in $\mathbb{R}^{n\times n}$.

(ii) For any compact subset $G \subset \mathbb{R}_+^{1\times n}$ and $D \in C(\mathbb{R}_+)^n$, it holds $GD \in C(\mathbb{R}_+)$.

It can be shown that $\phi(f|\mathcal{Q}) \in C(\mathbb{R}_+)^n$ in the following. The map $\mathcal{L} : C(\mathbb{R}_+)^n \to C(\mathbb{R}_+)^n$ is defined by

$$\mathcal{L}(f)\boldsymbol{v} = \boldsymbol{r}(f) + \beta\mathcal{Q}(f)\boldsymbol{v},\ \ \boldsymbol{v} \in C(\mathbb{R}_+)^n, \tag{5}$$

where $\mathcal{Q}(f) = \langle \underline{Q}(f), \overline{Q}(f)\rangle, \underline{Q}(f) = (\underline{q}_{ij}(f(i))) \in \mathbb{R}_+^{n\times n}, \overline{Q}(f) = (\overline{q}_{ij}(f(i))) \in \mathbb{R}_+^{n\times n}$.

From Lemma 1, we have $\mathcal{L}(f)\boldsymbol{v} \in C(\mathbb{R}_+)^n$ ($\boldsymbol{v} \in C(\mathbb{R}_+)^n$). Moreover, we define $\underline{L}(f) : \mathbb{R}_+^n \to \mathbb{R}_+^n, \overline{L}(f) : \mathbb{R}_+^n \to \mathbb{R}_+^n$ as follows: For $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)' \in \mathbb{R}_+^n$,

$$\underline{L}(f)\boldsymbol{x} = \boldsymbol{r}(f) + \beta \min_{Q \in \mathcal{Q}(f)} Q\boldsymbol{x}, \tag{6}$$

$$\overline{L}(f)\boldsymbol{x} = \boldsymbol{r}(f) + \beta \max_{Q \in \mathcal{Q}(f)} Q\boldsymbol{x}. \tag{7}$$

Then, we have the followings:

Lemma 2. *For any $f \in F$,*

(i) $\mathcal{L}(f)$ is monotone increasing and contractive mapping.

(ii) $\underline{L}(f)$ and $\overline{L}(f)$ are both monotone increasing and contractive mapping with respect to sup-*norm.*

Theorem 1. *For any $f \in F$, it holds that*

(i) $\phi(f|\mathcal{Q}) \in C(\mathbb{R}_+)^n$ and $\phi(f|\mathcal{Q})$ is the unique fixed point of $\mathcal{L}(f)$. Moreover, for any $\boldsymbol{v} \in C(\mathbb{R}_+)^n$, we have $\mathcal{L}(f)^\ell \boldsymbol{v} \to \phi(f|\mathcal{Q})$ ($\ell \to \infty$).

(ii) Let $\phi(f|\mathcal{Q}) = [\underline{\phi}(f), \overline{\phi}(f)]$. Then, $\underline{\phi}(f)$ and $\overline{\phi}(f)$ are the unique fixed point of $\underline{L}(f)$ and $\overline{L}(f)$, respectively.

Applying the result of De Robertis and Hartigan's (cf. [9]) on Bayesian inference method using intervals of prior measures, uncertain MDPs are formulated as interval estimated MDPs .

Let $P_n := P(S) = \{p = (p_1, p_2, \ldots, p_n) | p_i \geqq 0, \sum_{i=1}^n p_i = 1\}$ and $\mathcal{B}$ the set of all measurable set in $\mathbb{R}^n$, where $\mathbb{R}^n$ denotes the set of n-dimensional real vectors. For given measures L and U on $\mathcal{B}$, we denote by $L \leq U$ if $L(A) \leq U(A)$ for all set $A \in \mathcal{B}$. Let us denote by $[L, U]$ the convex set of measures Q satisfying $L(A) \leq Q(A) \leq U(A)$.

For simplicity, let the prior measures $[L, kL]$ $(k \geqq 1)$, where $L(\cdot)$ denotes a Lebesgue measure on P_n. We denote by $\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_n)$ a data of independent experiments, where the i-th state is observed with probability p_i and σ_i the number of outcomes of state i. Then, for a parameter $p = (p_1, p_2, \ldots, p_n) \in P_n$, a data set σ has probability density function of multinomial distribution as follows:

$$f(\sigma_1, \sigma_2, \ldots, \sigma_n | p) = \frac{(\sigma_1 + \cdots + \sigma_n)!}{\sigma_1! \cdots \sigma_n!} p_1^{\sigma_1} p_2^{\sigma_2} \cdots p_n^{\sigma_n}. \tag{8}$$

By using Bayesian inference for a prior measures $[L, U]$, intervals $[\underline{\lambda}_i, \overline{\lambda}_i]$ $(i \in S)$ of posterior measures of p_i is given by the range of integral ratios:

$$\left\{ \int_{P_n} p_i Q(dp) / \int_{P_n} Q(dp) \middle| L_\sigma \leqq Q \leqq U_\sigma \right\}, \tag{9}$$

where L_σ and U_σ are respectively lower and upper bounds of posterior measure for σ, and characterized as the unique solutions of (10) and (11):

Theorem 2. *Lower bound $\underline{\lambda}_i$ and upper bound $\overline{\lambda}_i$ of posterior intervals $[\underline{\lambda}_i, \overline{\lambda}_i]$ $(i \in S)$ are unique solutions as following equations:*

$$U_\sigma(p_i - \underline{\lambda}_i)^- + L_\sigma(p_i - \underline{\lambda}_i)^+ = 0, \tag{10}$$

$$U_\sigma(p_i - \overline{\lambda}_i)^+ + L_\sigma(p_i - \overline{\lambda}_i)^- = 0, \tag{11}$$

where $Q(f)$ denotes the integral of function f w.r.t. measure Q, $x^+ = \max\{0, x\}, x^- = x - x^+ = \min\{0, x\}$.

2 Bayesian control chart in uncertain MDPs

In this section, we describe a quality control model and formulate equivalent Bayesian model to the original problem.

Let $X(t)$ denote states of system at time $t (t \geqq 0)$ where $X(t) = \begin{cases} 0 & \text{(in-control)}, \\ 1 & \text{(out-of-control)}. \end{cases}$
The transition of state is occurred exponentially with distribution function $Exp(\theta)$, where $\Theta \ni \theta$ is unknown. We denote by $\widetilde{\theta}$ a random variable of θ. The partially observations of state $X(t)$ of system are made by sampling Y_i $(i = 1, 2, \ldots)$ of q-dimensional data of sample size n. The alternatives Δ are selected at each inspection step ih $(i = 1, 2, \ldots)$, where

$$\Delta = \begin{cases} 0 & \text{continue}, \\ 1 & \text{stop and search}. \end{cases}$$

It is noted that by the information of sampling date the decision maker choices between action "0" or "1" at each decision epoch $ih(i = 1, 2, \ldots)$.

We assume the following.

Assumption 1. If $X_{ih} = 0$ (or 1), $y_1^i, y_2^i, \ldots, y_n^i$ are i.i.d. random variables from $N_q(\mu_0, \Sigma)$ (or $N_q(\mu_1, \Sigma)$), where $N_q(\mu_0, \Sigma), N_q(\mu_1, \Sigma)$ are q-dimensional normal distribution and Σ is variance covariance matrix(positive definite) and μ_0, μ_1 are mean vector, respectively. For M-distance d_1 between μ_0 and μ_1 we assume the following:

$$d_1 := \left[(\mu_1 - \mu_0)\,\Sigma^{-1}\,(\mu_1 - \mu_0)^t\right]^{\frac{1}{2}} > 0. \tag{12}$$

The cost structures of this inspection model are given in the followings: Investigation cost $A > 0$ will be occurred to stop and research the system whether there is failure of system or not. When the state of system is failure, renewal cost $R \geqq 0$ will be charged to change the state from "1"(out-of-control) to "0"(in-control). If the system is operating without knowing the state of system being failure, operating cost $M > 0$ per unit time is incurred while remaining the system in out-of-control. If the decision maker selects an action "1"(stop and search), an observation cost $b + nc$ $(b, c \geqq 0)$ of sample size n is incurred.

Let sample space be $\overline{\Omega} = \Theta \times \Omega, \Omega = S \times (A \times S)^\infty$and random variables of process $\tilde{\theta}, \tilde{p}_0, \tilde{a}_0, \tilde{p}_1, \tilde{a}_1, \ldots$. That is, for $\omega \in \overline{\Omega} = (\theta, p_0, a_0, p_1, a_1, p_2, \ldots)$, we have $\tilde{\theta}(\omega) = \theta, \tilde{p}_0(\omega) = p_0, \tilde{a}_0(\omega) = a_0, \tilde{p}_1(\omega) = p_1, \ldots$, where we set $p_0 = 0$ without loss of generality. When the state is $\tilde{p}_m = p$ at epoch mh, an action $\tilde{a}_m = 0$(or $1, respectively$) is selected and at the next epoch $(m+1)h$ the information $Y_{m+1} = y^{m+1}$ is obtained and at epoch $(m+1)h$ the state moves to

$$\tilde{p}_{m+1} = T(p, y^{m+1}, 0)\left(\text{ or } T(p, y^{m+1}, 1), \text{ respectively}\right), \tag{13}$$

where, by Bayes' theorem, priori-posteriori Bayesian operator T is defined as in (Lemma 1 in [8]):

When θ is true parameter the costs $c(\cdot, \cdot)$ are given below:

$$\begin{cases} c(p, 0) = ME(\int_0^h I_{\{X_s=1\}}\, ds) + b + nc = M\left[h - \frac{1-p}{\theta}(1 - e^{-\theta h})\right] + b + nc, \\ c(p, 1) = c_1(p) + c(0, 0), \end{cases} \tag{14}$$

where $c_1(p) = A + Rp$.

Average expected cost $\varphi(\pi|\theta, p_0)$ given by $\tilde{\theta} = \theta \in \mathcal{P}(\Theta)$ and initial state distribution $p_0 = p \in S$ is defined as follows:

$$\varphi(\pi|\theta, p_0) = \limsup_{k \to \infty} \frac{1}{E(\tau_k)} E\left[\sum_{m=0}^{\tau_k} c(\tilde{p}_m, \tilde{a}_m)|\theta, p_0\right], \tag{15}$$

where, $\pi = (\tau_0, \tau_1, \tau_2, \ldots)$.

In addition, define discounted total expected cost $v(\pi|\theta, p_0)$ as follows:

$$v(\pi|\theta, p_0) = \sum_{m=0}^{\infty} \beta^m E_\pi \left[c_\alpha(\tilde{p}_m, \tilde{a}_m)|\theta, p_0\right], \tag{16}$$

where, $\beta = e^{-\alpha h}$ denotes discount rate and $E_\pi[\cdot|\theta, p]$ is expectation with probability measure $P_\pi(\cdot|\theta, p)$ on $\overline{\Omega}$ given by parameters θ, p and policy π.

Each policy $\pi \in \Pi$ which minimize $\varphi(\pi|\theta, p), v(\pi|\theta, p)$ respectively call θ-average optimal and θ-discounted optimal respectively. We have the following theorems.

Theorem 3 (V. Makis[8]). *If $A + R < \dfrac{M}{\theta}$, there exists θ-average optimal policy π *of the control-limit type. That is, there exists $p^*_\theta \in (0, 1)$ such that control policy following the decision function $f_\theta : S \to A$ as below is θ-average optimal.*

$$f_\theta(p) = \begin{cases} 0 & \text{if } p < p^*_\theta, \\ 1 & \text{if } p \geqq p^*_\theta. \end{cases} \tag{17}$$

Theorem 4. *(Sasaki, Horiguchi and Kurano ([10]))There exists θ-discounted optimal policy of control-limit type, that is, there exists $\overline{p}_\theta \in (0, 1)$ such that optimal decision function $g_\theta : S \to A$ is given as below:*

$$g_\theta(p) = \begin{cases} 0 & \text{if } p < \overline{p}_\theta, \\ 1 & \text{if } p \geqq \overline{p}_\theta. \end{cases} \tag{18}$$

3 Repair problem with exponentially deteriorating system

We consider the following repair problem. Let $\theta \in \Theta = \{\theta_1, \ldots, \theta_k\}$ denote a finite parameter space. For each $\theta_i \in \Theta$, we denote by $g(t|\theta)$ a pdf of deteriorating time t with exponential distribution as

$$g(t|\theta) = \begin{cases} \theta e^{-\theta t}, & t \geqq 0, \\ 0, & t < 0. \end{cases}$$

Let constant $T > 0$ be a time of inspection intervals. Let X_i denote the state of system and if $X_i = 0$ it means the system is under control and if $X_i = 1$ it means the system is out of control. If the state X_i is 0 at the inspection epoch iT, by the memorylessness property of exponential distribution, it does not affect the probability of the system being out of control from then on. On the other hand, if the state X_i is 1 at epoch iT, the system is repaired immediately and starts as new one after that.

It is easily seen that $P(X_i = 0) = P(X_i > T) = \int_T^\infty \theta e^{-\theta x} dx = e^{-\theta T}$ for each inspection epoch $iT, i = 1, 2, \ldots$, and $P(X_i = 1) = 1 - e^{-\theta T}$. Then we have a pdf of state of system given the parameter θ_i by $f(0|\theta_i) = e^{-\theta_i T}$, and $f(1|\theta_i) = 1 - e^{-\theta_i T}$.

In this inspection and repair problem, we consider the rate of convergence about posterior probabilities by the sequences of states of system inspections.

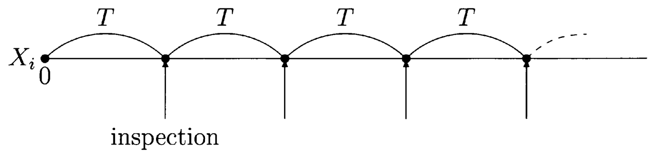

Figure 1: The states of system are inspected at each epoch.

Let $X_1, X_2, \ldots$ are i.i.d. random variables with $X_i \sim f(x|\theta)$ where $\theta \in \Theta = \{\theta_1, \theta_2, \ldots, \theta_k\}$. We write $q = (q(\theta_1), q(\theta_2), \ldots, q(\theta_k))$ for apriori distribution of parameter $\theta \in \Theta$. Let $h_n = (x_1, x_2, \ldots, x_n)$ denote the sample of size n where $X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n$, $x_i = 0$ or 1, $i = 1, 2, \cdots, n$. We will denote by $q_n(q, h_n)(\theta_i)$ the posteriori distribution of parameter $\theta \in \Theta$ as follows:

$$q_n(q, h_n)(\theta_i) = \frac{q(\theta_i)\prod_{l=1}^{n} f(x_l|\theta_i)}{\sum_{j=1}^{k} q(\theta_j)\prod_{l=1}^{n} f(x_l|\theta_j)}.$$

In order to show the rate of convergence of posteriori distribution, we give some well-known results.

Theorem 5 (Hölder's Inequality). *For $p > 1, \frac{1}{p} + \frac{1}{q} = 1, f \in L^p, g \in L^q$ and $\mu(\cdot)$ Lebesgue measure, it holds that*

$$\left| \int f(x)g(x)d\mu \right| \leqq \left(\int |f(x)|^p d\mu \right)^{\frac{1}{p}} \left(\int |g(x)|^q d\mu \right)^{\frac{1}{q}}. \tag{19}$$

Moreover, for $f'(x) := \frac{|f(x)|^p}{\int |f(x)|^p d\mu}, g'(x) := \frac{|g(x)|^q}{\int |g(x)|^q d\mu}$ and $D := \{x | f(x) \neq g(x)\}$, if $\mu(D) > 0$, it holds that $|\int f(x)g(x)d\mu| < ||f||_p \cdot ||g||_q$, i.e., the inequality (19) holds for strict inequality ($<$).

If $p = q = 2$ it is known as Schwartz's inequality and for probability density functions f and g for some distributions on probability space $(X, \mathcal{B}, P)$, we have the following.

Lemma 3. *Let $D := \{x | f(x) \neq g(x)\}$ and $D' := \{x | f'(x) \neq g'(x)\}$. For pdfs $f(x)$ and $g(x)$, if $\mu(D) > 0$ implies $\mu(D') > 0$.*

Proof. If $\mu(D') = 0$, i.e., $f'(x) = g'(x)$ a.s., then, there exist $c_1, c_2 > 0$ such that $\dfrac{f(x)^2}{c_1} = \dfrac{g(x)^2}{c_2}$. Hence there exists $c > 0$ such that $f(x) = cg(x)$. The integrals over x for each side function of equation, we have $1 = c \times 1$. Therefore we get $c = 1$. It means $f(x) = g(x)$ a.s , i.e., $\mu(D) = 0$, which proves the lemma by showing this contrapositive. ∎

Corollary 1. *(Schwartz's inequality for $\mu(D) > 0$) For pdfs f and g, if $\mu(D) > 0$ implies*

$$\left| \int (f(x))^{\frac{1}{2}} (g(x))^{\frac{1}{2}} d\mu \right| < 1.$$

We assume the following.

Assumption 2. $\mu(D_{ij}) > 0$ for $1 \leqq i < j \leqq k$, where $D_{ij} = \{x | f(x|\theta_i) \neq f(x|\theta_j)\}$.

Then, we have the following.

Theorem 6. *Under Assumption 2 and for a prior* $q = (q(\theta_1), q(\theta_2), \ldots, q(\theta_k))$ *with* $q(\theta_i) > 0$ *for all* $i (1 \leqq i \leqq k)$, *there exists* $\lambda (0 < \lambda < 1)$ *such that*

$$P(q_n(q, h_n)(\theta_i) > \varepsilon | \tilde{\theta} = \theta_{i_0}) \leqq K(q, n)\lambda^n$$

for any $\varepsilon > 0$ *and* $i \neq i_0$, *where* $K(q, n) = \max_{i \neq i_0} \frac{1}{\varepsilon^{\frac{1}{2}}} \left(\frac{q(\theta_i)}{q(\theta_{i_0})} \right)^{\frac{1}{2}}$.

Proof. Since $X_1, X_2, \ldots$ are i.i.d., we have

$$q_n(q, h_n)(\theta_i) = \frac{q(\theta_i) \prod_{l=1}^n f(x_l|\theta_i)}{\sum_{j=1}^k q(\theta_j) \prod_{l=1}^n f(x_l|\theta_j)} = \frac{q(\theta_i) \prod_{l=1}^n f(x_l|\theta_i)}{q(\theta_{i_0}) \prod_{l=1}^n f(x_l|\theta_{i_0})}.$$

Hence,

$$\begin{aligned}
&P\left(q_n(q, h_n)(\theta_i) > \varepsilon | \tilde{\theta} = \theta_{i_0}\right) \\
=&P\left(\frac{1}{\varepsilon} q_n(q, h_n)(\theta_i) > 1 | \tilde{\theta} = \theta_{i_0}\right) \\
=&P\left(\sqrt{\frac{1}{\varepsilon}} \sqrt{q_n(q, h_n)(\theta_i)} > 1 | \tilde{\theta} = \theta_{i_0}\right) \\
\leqq&E\left(\sqrt{\frac{1}{\varepsilon}} \sqrt{q_n(q, h_n)(\theta_i)} > 1 | \tilde{\theta} = \theta_{i_0}\right) \\
=&\frac{1}{\varepsilon^{\frac{1}{2}}} \int \sqrt{\frac{q(\theta_i) \prod_{l=1}^n f(x_l|\theta_i)}{q(\theta_{i_0}) \prod_{l=1}^n f(x_l|\theta_{i_0})}} \prod_{l=1}^n f(x_l|\theta_{i_0})\, dx_1 \cdots dx_n \\
=&\frac{q(\theta_i)^{\frac{1}{2}}}{\varepsilon^{\frac{1}{2}} \cdot q(\theta_{i_0})^{\frac{1}{2}}} \int \left(\prod_{l=1}^n f(x_l|\theta_i)\right)^{\frac{1}{2}} \left(\prod_{l=1}^n f(x_l|\theta_{i_0})\right)^{\frac{1}{2}} dx_1 \cdots dx_n \\
=&\frac{q(\theta_i)^{\frac{1}{2}}}{\varepsilon^{\frac{1}{2}} \cdot q(\theta_{i_0})^{\frac{1}{2}}} \prod_{l=1}^n \int f(x_l|\theta_i)^{\frac{1}{2}} f(x_l|\theta_{i_0})^{\frac{1}{2}} dx_l \quad \text{by Fubini's Theorem} \\
\leqq&\frac{1}{\varepsilon^{\frac{1}{2}}} \frac{q(\theta_i)^{\frac{1}{2}}}{q(\theta_{i_0})^{\frac{1}{2}}} \lambda^n = K(q, n)\lambda^n \quad \text{for } \exists \lambda (0 < \lambda < 1),
\end{aligned}$$

which completes the proof. ∎

Finally, we apply this theorem to our inspection and repair problem.

Let $\lambda_{i,i_0} = \int f(x|\theta)^{\frac{1}{2}} f(x|\theta_{i_0})^{\frac{1}{2}} d\mu$ for $i \neq i_0$. By Corollary 1, $\lambda_{i,i_0} < 1$. Let $\lambda = \max_{i \neq i_0} \lambda_{i,i_0}$. Then from Theorem 1, for $\tilde{\theta} = \theta_{i_0}$ we have

$$q_n(q, h_0)(\theta_k) \longrightarrow \begin{cases} 1 & (k = i), \\ 0 & (k \neq i_0) \end{cases}$$

with exponentially fast as $n \to \infty$.

References

[1] M. H. DeGroot. *Optimal statistical decisions.* McGraw-Hill Book Co., New York, 1970.

[2] T. S. Ferguson. *Mathematical Statistics.* Academic Press, New York - London. 1967.

[3] D. J. Hartfiel *Markov set-chains*, volume 1695 of *Lecture Notes in Mathematics.* Springer-Verlag, Berlin, 1998.

[4] O. Hernández-Lerma. *Adaptive Markov control processes*, volume 79 of *Applied Mathematical Sciences.* Springer-Verlag, New York, 1989.

[5] M. Kurano, J. Song, M. Hosaka and Y. Huang. Controlled Markov set-chains with discounting. *J. Appl. Probab.*, pages 293–302, 1998.

[6] M. Hosaka, M. Horiguchi and M. Kurano. Controlled Markov set-chains under average criteria. *Appl. Math. Comput.*, pages 195–209, 2001.

[7] M. Horiguchi. Bayesian Inference in Markov Decision Processes. In, *Modern Trends in Controlled Stochastic Processes: Theory and Applications, Vol. 2 (A.B. Piunovskiy ed.)*, Luniver Press (2015). (In press).

[8] Viliam Makis. Multivariate Bayesian control chart. *Oper. Res.*, 56(2):487–496, 2008.

[9] L. De Robertisand J. A. Hartigan, Bayesian inference using intervals of measures. *Ann. Statist.*, 9:235–244, 1981.

[10] M. Sasaki, M. Horiguchi and M. Kurano. Adaptive methods for multivariate bayesian control chart. *RIMS kokyuroku No. 1912 (In Japanese)*, pages 181–192, 2014.

[11] K. M. van Hee. *Bayesian control of Markov chains*, volume 95 of *Mathematical Centre Tracts.* Mathematisch Centrum, Amsterdam, 1978.

Masayuki Horiguchi
Department of Mathematics,
Faculty of Science, Kanagawa University
Address: Tsuchiya 2946, Hiratsuka City,
Kanagawa Prefecture, 259-1293, Japan
E-mail address: horiguchi@kanagawa-u.ac.jp

神奈川大学・理学部　堀口正之